UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A161 046

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: A.I. Memo 863

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): Direct Passive Navigation: Analytical Solution for Planes

5. TYPE OF REPORT & PERIOD COVERED: A.I. Memo - Sep. 84- July 85

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): Shahriar Negahdaripour

8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0643

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, VA 22209

12. REPORT DATE: August 85

13. NUMBER OF PAGES: 14

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, VA 22217

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Passive Navigation, Optical Flow, Structure & Motion, Planar Surfaces, Least Squares

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, we derive a closed form solution for recovering the motion of an observer relative to a planar surface directly from image brightness gradients. We do not compute the optical flow as an intermediate step, only the spatial and temporal intensity gradients at a minimum of eight points. ~~We solve~~ a linear matrix equation for the elements of a 3x3 matrix. The eigenvalue decomposition of its symmetric part is then used to compute the motion parameters and the plane orientation.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 863 August, 1985

Direct Passive Navigation: Analytical Solution for Planes

Shahriar Negahdaripour

Abstract: In this paper, we derive a closed form solution for recovering the motion of an observer relative to a planar surface directly from image brightness derivatives. We do not compute the optical flow as an intermediate step, only the spatial and temporal intensity gradients at a minimum of 8 points. We solve a linear matrix equation for the elements of a 3x3 matrix. The eigenvalue decomposition of its symmetric part is then used to compute the motion parameters and the plane orientation.

Key Words: Passive Navigation, Optical Flow, Structure and Motion, Planar Surfaces, Least Squares.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Avdanced Research Projects Agency of the Department of Defense under office of Naval Research contract N00014-75-C-0643 and in part by System Development Foundation.

1. Introduction

The problem of determining rigid body motion and surface structure from image data has been the topic of many research papers in the area of machine vision [1-22]. Many approaches based on, tracking feature points [5,11,19,20] or contours [9], using motion flow field [1,3,4,10,12,16,17,21,22], texture [2], or image intensity gradients [14,15] have been proposed in the literature.

In the feature point matching schemes, information about a finite number of well-seperated points is used to recover the motion (general 8-point 2-frame algorithms of Longuet-Higgins [11], Tsai and Huang [20], Buxton et al. [5], and the algorithm of Tsai, Huang and Zhu [19] for planar surfaces). These methods require identifying and matching feature points in a sequence of images. The minimum number of points required depends on the number of image frames. With 2 frames, in most cases, a minimum of 5 points results in a unique solution from a set of nonlinear equations. However, using 8 points, as in algorithms cited above, one only solves linear equations. Here, it is assumed that the more difficult problem of establishing point correspondence has already been solved. In general, this involves determining corners along contours using iterative searches. For images of smooth objects, it is difficult to find good features or corners.

For the general case of smooth surfaces, Longuet-Higgins and Prazdny [11] suggested a method that uses the optical flow and its first and second derivatives at a single point. Later, Waxman and Ullman [21] developed this into an algorithm for recovering the structure and motion parameters from a set of nonlinear equations. Subbarao and Waxman [17] recently found a closed form solution to the original formulation in [21] for planar surfaces. These methods while mathematically elegant are very sensitive to errors in the optical flow data since second order derivatives of noisy data are used.

At the expense of more computation, more robust algorithms have been suggested using the optical flow at every image point [1,3,4].

Longuet-Higgins [12] has presented a closed form solution for planar surfaces, very similar to ours, using the coefficients of the second order optical flow equations. However, it is assumed that both components of the flow field have already been computed for a minimum of 5 image points.

By representing a planar surface in the form of a closed contour, Kanatani [9] has shown that the surface and motion parameters can be computed by measuring "diameters" of the contour using line and surface integrals. Here, no point correspondence is required.

Assuming that the planar surface has a uniform texture density, Aloimonos and Chou [2] have presented a procedure for computing the motion and surface orientation from texture.

In much of the research work in recovering surface structure and motion from the optical flow field, it is assumed that a reasonable estimate of the full optical flow field is available. In general, the computation of the local flow field exploits a constraint equation between the local intensity changes and the two components of the optical flow. However, this only gives the component of the flow in the direction of the intensity gradient. To compute the full flow field, one needs additional constraints such as the

heuristic assumption that the flow field is locally smooth [7,8]. This, in many cases, leads to optical flow fields that are not consistant with the true motion field.

In an earlier paper, we presented an iterative scheme for recovering the motion of an observer relative to a planar surface directly from the image brightness derivatives, and without the need to compute the local flow field [14,15]. Further, using a compact vector notation, we showed that, at most, two interpretations are possible for planar surfaces and derived the relationship between them. Here, we present a closed form solution to the same problem. We first solve a linear matrix equation for the elements of a 3x3 matrix using intensity derivatives at a minimum of 8 non-colinear points. The special structure of this matrix allows us to compute the motion and structure parameters very easily.

2. Preliminaries

We first recall some details about perspective projection, the motion field, the brightness change constraint equation, rigid body motion and planar surfaces. This we do using vector notation in order to keep the resulting equations as compact as possible.

2.1. Perspective Projection

Let the center of projection be at the origin of a Cartesian coordinate system. Without loss of generality we assume that the effective focal length is unity. The image is formed on the plane $z = 1$, parallel to the xy-plane, that is, the optical axis lies along the z-axis. Let $\mathbf{R}$ be a point in the scene. Its projection in the image is $\mathbf{r}$, where

$$\mathbf{r} = \frac{1}{\mathbf{R}\cdot\hat{\mathbf{z}}}\mathbf{R}.$$

The z-component of $\mathbf{r}$ is clearly equal to one, that is $\mathbf{r}\cdot\hat{\mathbf{z}} = 1$.

2.2. Motion Field and Optical Flow

The *motion field* is the vector field induced in the image plane by the relative motion of the observer with respect to the environment. The *optical flow* is the apparent motion of brightness patterns. Under favourable circumstances the optical flow is identical to the motion field (moving shadows or uniform objects in motion could create descrepencies between the motion field and the optical flow. Here, we assume that the motion flow field and the optical flow are the same). The velocity of the image $\mathbf{r}$ of a point $\mathbf{R}$ is given by

$$\frac{d\mathbf{r}}{dt} = \frac{d}{dt}\frac{1}{\mathbf{R}\cdot\hat{\mathbf{z}}}\mathbf{R}.$$

For convenience, we introduce the notation $\mathbf{r}_t$ and $\mathbf{R}_t$ for the time derivatives of $\mathbf{r}$ and $\mathbf{R}$, respectively. We then have

$$\mathbf{r}_t = \frac{1}{\mathbf{R}\cdot\hat{\mathbf{z}}}\mathbf{R}_t - \frac{1}{(\mathbf{R}\cdot\hat{\mathbf{z}})^2}(\mathbf{R}_t\cdot\hat{\mathbf{z}})\mathbf{R},$$

which can also be written in the compact form

$$\mathbf{r}_t = \frac{1}{(\mathbf{R}\cdot\hat{\mathbf{z}})^2}(\hat{\mathbf{z}} \times (\mathbf{R}_t \times \mathbf{R})),$$

since $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{c}\cdot\mathbf{a})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$. The vector $\mathbf{r}_t$ lies in the image plane, and so $(\mathbf{r}_t \cdot \hat{\mathbf{z}}) = 0$. Further, $\mathbf{r}_t = \mathbf{0}$, if $\mathbf{R}_t \parallel \mathbf{R}$, as expected.

Finally, noting that $\mathbf{R} = (\mathbf{R}\cdot\hat{\mathbf{z}})\mathbf{r}$, we get

$$\mathbf{r}_t = \frac{1}{\mathbf{R}\cdot\hat{\mathbf{z}}}(\hat{\mathbf{z}} \times (\mathbf{R}_t \times \mathbf{r})).$$

2.3. Rigid Body Motion

In the case of the observer moving relative to a rigid environment with translational velocity $\mathbf{t}$ and rotational velocity ω, we find that the motion of a point in the environment relative to the observer is given by

$$\mathbf{R}_t = -\omega \times \mathbf{R} - \mathbf{t}.$$

Since $\mathbf{R} = (\mathbf{R}\cdot\hat{\mathbf{z}})\mathbf{r}$, we can write this as

$$\mathbf{R}_t = -(\mathbf{R}\cdot\hat{\mathbf{z}})\,\omega \times \mathbf{r} - \mathbf{t}.$$

Substituting for $\mathbf{R}_t$ in the formula derived above for $\mathbf{r}_t$, we obtain

$$\mathbf{r}_t = -\Big(\hat{\mathbf{z}} \times (\mathbf{r} \times (\mathbf{r} \times \omega - \frac{1}{\mathbf{R}\cdot\hat{\mathbf{z}}}\mathbf{t}))\Big).$$

It is important to remember that there is an inherent ambiguity here, since the same motion field results when distance and the translational velocity are multiplied by an arbitrary constant. This can be seen easily from the above equation since the same image plane velocity is obtained if one multiplies both $\mathbf{R}$ and $\mathbf{t}$ by some constant.

2.4. Brightness Change Equation

The brightness of the image of a particular patch of a surface depends on many factors. It may for example vary with the orientation of the patch. In many cases, however, it remains at least approximately constant as the surface moves in the environment. If we assume that the image brightness of a patch remains constant, we have

$$\frac{dE}{dt} = 0,$$

or

$$\frac{\partial E}{\partial \mathbf{r}} \cdot \frac{d\mathbf{r}}{dt} + \frac{\partial E}{\partial t} = 0,$$

where $\partial E/\partial \mathbf{r} = (\partial E/\partial x, \partial E/\partial y, 0)^T$ is the image brightness gradient. It is convenient to use the notation $E_{\mathbf{r}}$ for this quantity and E_t for the time derivative of the brightness.

Then we can write the brightness change equation in the simple form

$$E_{\mathbf{r}} \cdot \mathbf{r}_t + E_t = 0.$$

Substituting for $\mathbf{r}_t$ we get

$$E_t - E_{\mathbf{r}} \cdot \left(\hat{\mathbf{z}} \times (\mathbf{r} \times (\mathbf{r} \times \omega - \frac{1}{\mathbf{R} \cdot \hat{\mathbf{z}}}\mathbf{t}))\right) = 0.$$

Now

$$E_{\mathbf{r}} \cdot (\hat{\mathbf{z}} \times (\mathbf{r} \times \mathbf{t})) = (E_{\mathbf{r}} \times \hat{\mathbf{z}}) \cdot (\mathbf{r} \times \mathbf{t}) = \left((E_{\mathbf{r}} \times \hat{\mathbf{z}}) \times \mathbf{r}\right) \cdot \mathbf{t},$$

and by similar reasoning

$$E_{\mathbf{r}} \cdot \left(\hat{\mathbf{z}} \times (\mathbf{r} \times (\mathbf{r} \times \omega))\right) = \left(((\mathbf{E}_{\mathbf{r}} \times \hat{\mathbf{z}}) \times \mathbf{r}) \times \mathbf{r}\right) \cdot \omega,$$

so we have

$$E_t - \left(((E_{\mathbf{r}} \times \hat{\mathbf{z}}) \times \mathbf{r}) \times \mathbf{r}\right) \cdot \omega + \frac{1}{\mathbf{R} \cdot \hat{\mathbf{z}}}((\mathbf{E}_{\mathbf{r}} \times \hat{\mathbf{z}}) \times \mathbf{r}) \cdot \mathbf{t} = 0.$$

To make this constraint equation more compact, let us define $c = E_t$, $\mathbf{s} = (E_{\mathbf{r}} \times \hat{\mathbf{z}}) \times \mathbf{r}$, and $\mathbf{v} = -\mathbf{s} \times \mathbf{r}$; then, finally,

$$c + \mathbf{v} \cdot \omega + \frac{1}{\mathbf{R} \cdot \hat{\mathbf{z}}}\mathbf{s} \cdot \mathbf{t} = 0.$$

This is the brightness change equation in the case of rigid body motion.

2.5. Planar Surface

A particularly impoverished scene is one consisting of a single planar surface. The equation for such a surface is

$$\mathbf{R} \cdot \mathbf{n} = 1,$$

where $\mathbf{n}/|\mathbf{n}|$ is a unit normal to the plane, and $1/|\mathbf{n}|$ is the perpendicular distance of the plane from the origin. Since $\mathbf{R} = (\mathbf{R} \cdot \hat{\mathbf{z}})\mathbf{r}$, we can write this as

$$\mathbf{r} \cdot \mathbf{n} = \frac{1}{\mathbf{R} \cdot \hat{\mathbf{z}}},$$

so the constraint equation becomes

$$c + \mathbf{v} \cdot \omega + (\mathbf{r} \cdot \mathbf{n})(\mathbf{s} \cdot \mathbf{t}) = 0.$$

This is the brightness change equation for a planar surface. Note again the inherent ambiguity in the constraint equation. It is satisfied equally well by two planes with the same orientation but at different distances provided that the translational velocities are in the same proportions.

2.6. Essential Parameters for Planar Surfaces

The brightness change equation can be written as:

$$c + (\mathbf{r} \times \mathbf{s}) \cdot \omega + (\mathbf{r} \cdot \mathbf{n})(\mathbf{s} \cdot \mathbf{t}) = 0.$$

Now, using the identity $(\mathbf{r} \times \mathbf{s}) \cdot \omega = (\mathbf{s} \times \omega) \cdot \mathbf{r}$, we get:

$$c + (\mathbf{s} \times \omega) \cdot \mathbf{r} + (\mathbf{r} \cdot \mathbf{n})(\mathbf{s} \cdot \mathbf{t}) = 0.$$

Let us define:

$$\mathbf{W} = \begin{pmatrix} 0 & -\omega_3 & +\omega_2 \\ +\omega_3 & 0 & -\omega_1 \\ -\omega_2 & +\omega_1 & 0 \end{pmatrix},$$

then, $(\mathbf{s} \times \omega) = \mathbf{W}\mathbf{s}$, in which case, we arrive at:

$$c + \mathbf{r}^T\mathbf{W}\mathbf{s} + \mathbf{r}^T\mathbf{n}\mathbf{t}^T\mathbf{s} = 0,$$

and after collecting terms:

$$c + \mathbf{r}^T(\mathbf{W} + \mathbf{n}\mathbf{t}^T)\mathbf{s} = 0.$$

Now let:

$$\mathbf{P} = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \\ p_7 & p_8 & p_9 \end{pmatrix} = \mathbf{W} + \mathbf{n}\mathbf{t}^T,$$

We will refer to p_i , $i = 1, 2, \ldots, 9$ as the *essential parameters* (in agreement with Tsai and Huang [20]) since these parameters contain all the information about the planar surface orientation and motion parameters. Substituting $\mathbf{P}$ into the brightness change equation, we get:

$$c + \mathbf{r}^T\mathbf{P}\mathbf{s} = 0.$$

This is a linear constraint equation in terms of the elements of $\mathbf{P}$, and can be used to solve for these parameters. We will show how the special structure of $\mathbf{P}$ can be exploited to recover the motion and structure parameters very easily.

3. Recovering Essential Parameters

The nine essential parameters satisfy the following constraint equation:

$$c + \mathbf{r}^T\mathbf{P}\mathbf{s} = 0.$$

or in terms of $\mathbf{p}$, a vector of length 9 whose i-th element is p_i:

$$c + \mathbf{a}^T\mathbf{p} = 0,$$

where:

$$\mathbf{a} = (r_1s_1, r_1s_2, r_1s_3, r_2s_1, r_2s_2, r_2s_3, r_3s_1, r_3s_2, r_3s_3)^T.$$

We can compute them using image brightness $E(x, y, t)$, and its spatial and time derivatives, $E_\mathbf{r}$ and E_t, over some region I in the image plane. Since there are only eight motion

and surface parameters to recover (There are three components of each of ω, $\mathbf{t}$, and $\mathbf{n}$. However, as mentioned earlier, the translational velocity and the surface normal can be recovered up to a scale factor.), only eight of the p_i's are independent. This implies that we can arbitrarily fix one of the essential parameters, and compute the remaining eight using eight independent points (At each point, we get one constraint and we have eight unknowns to recover).

Let $\mathbf{p}' = (p'_1, p'_2, ..., p'_8, 0)$ denote the solution obtained by setting the last element equal to zero. If we define:

$$\tilde{\mathbf{p}}' = (p'_1, p'_2, ..., p'_8), \quad \text{and} \quad \tilde{\mathbf{a}} = (r_1s_1, r_1s_2, r_1s_3, r_2s_1, r_2s_2, r_2s_3, r_3s_1, r_3s_2)^T,$$

then the above constraint equation reduces to:

$$\tilde{\mathbf{a}}^T \tilde{\mathbf{p}}' + c = 0.$$

Using eight non-colinear points (this is necessary to guarantee that the resulting equations are independent), we can solve the following linear matrix equation:

$$\mathbf{A}\tilde{\mathbf{p}}' + \mathbf{c} = 0,$$

where:

$$\mathbf{A} = (\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, ..., \tilde{\mathbf{a}}_8)^T, \quad \mathbf{c} = (c_1, c_2, ..., c_8).$$

The solution of the above equation is:

$$\tilde{\mathbf{p}}' = -\mathbf{A}^{-1}\mathbf{c}.$$

Image brightness values are distorted with sensor noise and quantization error. These inaccuracies are further accentuated by methods used for estimating the brightness gradient. Thus it is not advisable to base a method on measurements at just a few points. Instead we propose to minimize the error in the brightness constraint equation over the whole region I in the image plane. So we choose the essential parameters that minimize:

$$\iint_I (\tilde{\mathbf{a}}^T \tilde{\mathbf{p}}' + c)^2 \, dx \, dy.$$

The solution, in this case, is given by:

$$\tilde{\mathbf{p}}' = -\left(\iint_I \tilde{\mathbf{a}}\tilde{\mathbf{a}}^T \, dx \, dy\right)^{-1} \left(\iint_I c\tilde{\mathbf{a}} \, dx \, dy\right).$$

Note that, in general, the true p_9 is nonzero. We can show that the solution obtained through the assumption that $p'_9 = 0$, $\mathbf{p}'$, and the true solution (denoted by $\mathbf{p}$) are related by the equation:

$$\mathbf{p} = \mathbf{p}' + p_9\mathbf{u}, \quad \mathbf{u} = (1, 0, 0, 0, 1, 0, 0, 0, 1)^T.$$

The proof goes as follows. Since $\mathbf{s} = (E_\mathbf{r} \times \hat{\mathbf{z}}) \times \mathbf{r}$, then $\mathbf{r}^T\mathbf{s} = \mathbf{r} \cdot ((E_\mathbf{r} \times \hat{\mathbf{z}}) \times \mathbf{r}) = 0$. For any arbitrary constant l, such that $\mathbf{L} = l\mathbf{I}$ ($\mathbf{I}$ is the identity matrix), we have:

$$\mathbf{r}^T\mathbf{L}\mathbf{s} = 0.$$

If we add this to our costraint equation:

$$c + \mathbf{r}^T(\mathbf{W} + \mathbf{n}\mathbf{t}^T + \mathbf{L})\mathbf{s} = 0.$$

It is immediately apparent that any $\mathbf{P}'$ of the form:

$$\mathbf{P}' = \begin{pmatrix} p_1' & p_2' & p_3' \\ p_4' & p_5' & p_6' \\ p_7' & p_8' & p_9' \end{pmatrix} = \mathbf{W} + \mathbf{n}\mathbf{t}^T + \mathbf{L}$$

will also satisfy our constraint equation. Therefore, the two solutions for the essential parameters are related by:

$$\mathbf{P} = \mathbf{P}' - \mathbf{L},$$

or in terms of $\mathbf{p}$ and $\mathbf{p}'$:

$$\mathbf{p} = \mathbf{p}' - l\mathbf{u}, \quad \mathbf{u} = (1,0,0,0,1,0,0,0,1),$$

for some constant l.It follows that $p_9 = p_9' - l$. Now if $p_9' = 0$, then $l = -p_9$, and so:

$$\mathbf{p} = \mathbf{p}' + p_9\mathbf{u}.$$

In term of matrices $\mathbf{P}$, and $\mathbf{P}'$:

$$\mathbf{P} = \mathbf{P}' + p_9\mathbf{I}.$$

The procedure for determining p_9, and subsequently $\mathbf{P}$ is presented in the appendix. For now, we assume that $\mathbf{P}$ is known (note that we have, so far, shown how to compute $\mathbf{P}'$ and not $\mathbf{P}$).

4. Recovering Motion and Structure

We now show how to compute the parameters of the translational motion and the plane orientation from the essential parameters. Once these are known, the rotational parameters are determined from:

$$\mathbf{W} = \mathbf{P} - \mathbf{n}\mathbf{t}^T$$

Since $\mathbf{W}$ is skew-symmetric, and from the definition of $\mathbf{P}$:

$$\mathbf{P}^* = \mathbf{P} + \mathbf{P}^T = (W + W^T) + (\mathbf{t}\mathbf{n}^T + \mathbf{n}\mathbf{t}^T) = (\mathbf{t}\mathbf{n}^T + \mathbf{n}\mathbf{t}^T).$$

Consider the special case where $\mathbf{t} \parallel \mathbf{n}$ ($\mathbf{n} = k\mathbf{t}$, for some constant k). In this case,

$$\mathbf{P}^* = 2k\mathbf{t}\mathbf{t}^T$$

Computing $\mathbf{t}$ from the above equation, and subsequently $\mathbf{n}$ is a simple practice of algebra. Therefore, from now, we assume that $\mathbf{t} \times \mathbf{n} \neq \mathbf{0}$.

Since the translational vector and the surface normal can be recovered up to a scale factor, we can, without loss of generality, let $\mathbf{t}^T\mathbf{t} = 1$. We will present our results in terms of the normalized $\mathbf{n}$ and $\mathbf{t}$, that is:

$$\tilde{\mathbf{n}} = \frac{\mathbf{n}}{|\mathbf{n}|}, \quad \tilde{\mathbf{t}} = \frac{\mathbf{t}}{|\mathbf{t}|} = \mathbf{t}.$$

Note that:

$$\text{Trace}\mathbf{P}^* = (\mathbf{n}^T\mathbf{t} + \mathbf{t}^T\mathbf{n}) = 2(\mathbf{n}\cdot\mathbf{t}) = 2\,|\mathbf{n}|\,(\tilde{\mathbf{n}}\cdot\tilde{\mathbf{t}}),$$

and therefore, we can always copmute $\mathbf{n}$ from:

$$\mathbf{n} = |\mathbf{n}|\,\tilde{\mathbf{n}}, \quad |\mathbf{n}| = \frac{\text{Trace}\mathbf{P}^*}{2(\tilde{\mathbf{n}}\cdot\tilde{\mathbf{t}})}$$

We also have to normalize $\mathbf{P}^*$. From now, $\mathbf{P}^*$ denotes the normalized matrix, so that:

$$\mathbf{P}^* = (\tilde{\mathbf{t}}\tilde{\mathbf{n}}^T + \tilde{\mathbf{n}}\tilde{\mathbf{t}}^T).$$

In order to present the solutions for $\tilde{\mathbf{n}}$ and $\tilde{\mathbf{t}}$, it is necessary to express the eigenvalue decomposition of $\mathbf{P}^*$ in terms of these vectors. We will do so in the next lemma.

Lemma 1: Let $\mathbf{P}^* = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ be the eigenvalue decomposition of $\mathbf{P}^* = (\tilde{\mathbf{t}}\tilde{\mathbf{n}}^T + \tilde{\mathbf{n}}\tilde{\mathbf{t}}^T)$, and let $\tau = \frac{1}{2}\text{Trace}\mathbf{P}^*$. Then:

$$\mathbf{\Lambda} = \begin{pmatrix} \tau-1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \tau+1 \end{pmatrix},$$

$$\mathbf{U} = \left[\frac{\tilde{\mathbf{t}}-\tilde{\mathbf{n}}}{\sqrt{2(1-\tau))}} \; \frac{\tilde{\mathbf{t}}\times\tilde{\mathbf{n}}}{|\tilde{\mathbf{t}}\times\tilde{\mathbf{n}}|} \; \frac{\tilde{\mathbf{t}}+\tilde{\mathbf{n}}}{\sqrt{2(1+\tau))}}\right].$$

Proof: $(\tilde{\mathbf{t}}\tilde{\mathbf{n}}^T + \tilde{\mathbf{n}}\tilde{\mathbf{t}}^T)\tilde{\mathbf{t}}\times\tilde{\mathbf{n}} = \tilde{\mathbf{t}}\big(\tilde{\mathbf{n}}\cdot(\tilde{\mathbf{t}}\times\tilde{\mathbf{n}})\big) + \tilde{\mathbf{n}}\big(\tilde{\mathbf{t}}\cdot(\tilde{\mathbf{t}}\times\tilde{\mathbf{n}})\big) = 0 = 0\cdot(\tilde{\mathbf{t}}\times\tilde{\mathbf{n}})$, and so $(\tilde{\mathbf{t}}\times\tilde{\mathbf{n}})$ is the eigenvector corresponding to the zero eigenvalue of $\mathbf{P}^*$. Since it is symmetric, $\mathbf{P}^*$ has 3 orthogonal eigenvectors. The other two eigenvectors are, therefore, in the plane containing $\tilde{\mathbf{t}}$ and $\tilde{\mathbf{n}}$. Let $\mathbf{u} = \alpha\tilde{\mathbf{t}} + \beta\tilde{\mathbf{n}}$ and λ denote an eigenvector-eigenvalue pair for some α and β (to be determined). Then

$$(\tilde{\mathbf{t}}\tilde{\mathbf{n}}^T + \tilde{\mathbf{n}}\tilde{\mathbf{t}}^T)(\alpha\tilde{\mathbf{t}} + \beta\tilde{\mathbf{n}}) = \lambda(\alpha\tilde{\mathbf{t}} + \beta\tilde{\mathbf{n}}),$$

which simplifies to

$$[\alpha(\tilde{\mathbf{t}}\cdot\tilde{\mathbf{n}}) + \beta(\tilde{\mathbf{n}}\cdot\tilde{\mathbf{n}})]\tilde{\mathbf{t}} + [\alpha(\tilde{\mathbf{t}}\cdot\tilde{\mathbf{t}}) + \beta(\tilde{\mathbf{t}}\cdot\tilde{\mathbf{n}})]\tilde{\mathbf{n}} = \lambda\alpha\tilde{\mathbf{t}} + \lambda\beta\tilde{\mathbf{n}}.$$

Since $(\tilde{\mathbf{t}}\cdot\tilde{\mathbf{n}}) = \tau$, we have:

$$\begin{pmatrix} \tau-\lambda & 1 \\ 1 & \tau-\lambda \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \mathbf{0}.$$

The solutions for λ are given by:

$$\lambda = \tau \pm 1.$$

Substituting for λ into the earlier equations, we get:

$$\alpha = \pm\beta.$$

Using these into the equation for $\mathbf{u}$, and normalizing the results yield:

$$\mathbf{u} = \frac{\tilde{\mathbf{t}} \pm \tilde{\mathbf{n}}}{\sqrt{2(1 \pm \tau)}}.$$

Note that since $|\tau| < 1$, $\lambda_1 < \lambda_2 = 0 < \lambda_3$. ∎

We can now detrmine $\tilde{\mathbf{t}}$ and $\tilde{\mathbf{n}}$. Let $\mathbf{u}_i$ denote the i-th column of $\mathbf{U}$. From lemma 1:

$$\mathbf{u}_1 = \frac{\tilde{\mathbf{t}} - \tilde{\mathbf{n}}}{\sqrt{2(1 - \tau)}}, \quad \mathbf{u}_3 = \frac{\tilde{\mathbf{t}} + \tilde{\mathbf{n}}}{\sqrt{2(1 + \tau)}}.$$

From the expressions for the eigenvectors, it follows that:

$$\tilde{\mathbf{t}} + \tilde{\mathbf{n}} = \sqrt{2(1+\tau)}\mathbf{u}_3, \quad \tilde{\mathbf{t}} - \tilde{\mathbf{n}} = \sqrt{2(1-\tau)}\mathbf{u}_1.$$

Solving these for $\tilde{\mathbf{t}}$ and $\tilde{\mathbf{n}}$:

$$\tilde{\mathbf{n}} = \sqrt{\frac{1+\tau}{2}}\mathbf{u}_3 - \sqrt{\frac{1-\tau}{2}}\mathbf{u}_1, \quad \tilde{\mathbf{t}} = \sqrt{\frac{1+\tau}{2}}\mathbf{u}_3 + \sqrt{\frac{1-\tau}{2}}\mathbf{u}_1,$$

Since the choice of the signs of the eigenvectors are arbitrary, we should repeat the above procedure with the sign of either $\mathbf{u}_1$ or $\mathbf{u}_3$ reversed:

$$\mathbf{u}_1 = -\frac{\tilde{\mathbf{t}} - \tilde{\mathbf{n}}}{\sqrt{2(1 - \tau)}}, \quad \mathbf{u}_3 = \frac{\tilde{\mathbf{t}} + \tilde{\mathbf{n}}}{\sqrt{2(1 + \tau)}}.$$

In this case, the second solution is obtained from the following equations:

$$\tilde{\mathbf{n}} = \sqrt{\frac{1+\tau}{2}}\mathbf{u}_3 + \sqrt{\frac{1-\tau}{2}}\mathbf{u}_1, \quad \tilde{\mathbf{t}} = \sqrt{\frac{1+\tau}{2}}\mathbf{u}_3 - \sqrt{\frac{1-\tau}{2}}\mathbf{u}_1,$$

Note that this is the dual solution for the planar surfaces. The special case of $\tilde{\mathbf{t}} \parallel \tilde{\mathbf{n}}$ corresponds to when the matrix $\mathbf{P}^*$ has multiple eigenvalues. Then, either:

1) $\tau = 1$, for which $\lambda_1 = \lambda_2 = 0$. Then the two solutions merge to the single one:

$$\tilde{\mathbf{n}} = \tilde{\mathbf{t}} = \mathbf{u}_3,$$

or:

2) $\tau = -1$, so that $\lambda_2 = \lambda_3 = 0$. Here, the unique solution is given by:

$$\tilde{\mathbf{n}} = -\tilde{\mathbf{t}} = \mathbf{u}_1.$$

As mentioned earlier, we can determine $\mathbf{n}$ through proper scaling of $\tilde{\mathbf{n}}$, and solve for the rotation parameters by substituting the solutions for $\mathbf{n}$ and $\mathbf{t}$ into the equation:

$$\mathbf{W} = \mathbf{P} - \mathbf{n}\mathbf{t}^T.$$

Even though we gave a complete and compact proof of the dual solution in an earlier paper[15], it is intriguing to confirm those results with our closed form solution. We showed that the two solutions are related by:

$$\mathbf{n}^* = k\mathbf{t}, \quad \mathbf{t}^* = 1/k\,\mathbf{n}, \quad \omega^* = \omega + \mathbf{n} \times \mathbf{t},$$

where k is arbitrarily chosen to scale the translation vector and plane normal. The two solutions given in lemma 2 for $\mathbf{n}$ and $\mathbf{t}$ already satisfy the duality relationship given above. It remains to show the same for the rotation parameters. We only have to show that:

$$\mathbf{W}^* + \mathbf{n}^*\mathbf{t}^{*T} = \mathbf{P},$$

where

$$\mathbf{W}^* = \mathbf{W} + \begin{pmatrix} 0 & n_1t_2 - n_2t_1 & n_1t_3 - n_3t_1 \\ n_2t_1 - n_1t_2 & 0 & n_2t_3 - n_3t_2 \\ n_3t_1 - n_1t_3 & n_3t_2 - n_2t_3 & 0 \end{pmatrix},$$

or

$$\mathbf{W}^* = \mathbf{W} + \mathbf{n}\mathbf{t}^T - \mathbf{t}\mathbf{n}^T.$$

Substituting for $\mathbf{n}^*$, $\mathbf{t}^*$, and $\mathbf{W}^*$ into the earlier equation, and simplifying the results, we get:

$$\mathbf{W} + \mathbf{n}\mathbf{t}^T = \mathbf{P}.$$

5. Summary

In this paper, we presented a closed form solution for recovering the motion of an observer with respect to a planar surface without having to compute the optical flow as an intermediate step. We only need the image intensity gradients at 8 non-colinear points, however, in general, it is better to compute these gradients in a larger portion of the image to reduce the noise effects. We first employed a constraint equation we developed for planar surfaces to compute 9 intermediate parameters, the elements of a 3x3 matrix. We referred to them as *essential parameters.* The special structure of this matrix allows us to compute the motion and plane parameters very easily.

6. Acknowledgements

The author would like to thank Muralidhara Subbarao for encouraging the investigation of closed form solution for planar surfaces, and Berthold K.P. Horn, and Alan Yuille for their helpful comments.

Appendix

In the previous sections, we showed how the motion paprameters can be recovered once the essential paprameters are known. However, the brightness change constraint equation allowed us to determine the matrix $\mathbf{P}'$ (a particular solution of $\mathbf{P}$ with the last element set to zero). We showed that the two solution are related by:

$$\mathbf{P} = \mathbf{P}' + p_9\mathbf{I}.$$

Here, we show how to determine p_9, and consequently $\mathbf{P}$. For simplicity, let $p_9 = l$, and let $\mathbf{P} = \mathbf{U}\Lambda\mathbf{V}^T$ denote the eigenvalue decomposition of $\mathbf{P}$, where $(\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3))$, then:

$$\mathbf{P}' = \mathbf{U}\Lambda\mathbf{V}^T - l\mathbf{I} = \mathbf{U}\Lambda\mathbf{V}^T - l\mathbf{U}\mathbf{V}^T.$$

If $\mathbf{L} = l\mathbf{I}$ then:

$$\mathbf{P}' = \mathbf{U}\Lambda\mathbf{V}^T - \mathbf{U}\mathbf{L}\mathbf{V}^T = \mathbf{U}(\Lambda - \mathbf{L})\mathbf{V}^T.$$

Similarly, if $\mathbf{P}^* = \mathbf{P} + \mathbf{P}^T = \mathbf{U}\Lambda\mathbf{U}^T$ denotes the eigenvalue decomposition of $\mathbf{P}^*$, then:

$$\mathbf{P}'^* = \mathbf{P}' + \mathbf{P}'^T = \mathbf{P} + \mathbf{P}^T - 2\mathbf{L} = \mathbf{P}^* - 2\mathbf{L} = \mathbf{U}(\Lambda - 2\mathbf{L})\mathbf{U}^T.$$

In lemma 1, we showed that $\lambda_1 < \lambda_2 = 0 < \lambda_3$ (and when $\mathbf{t} \parallel \mathbf{n}$, we get two zero eigenvalues). Therefore, the eigenvalues of $\mathbf{P}'^*$ can be arranged in the form:

$$\lambda_1 - 2l < -2l < \lambda_3 - 2l.$$

It follows that $l = -\frac{1}{2}\lambda_2(\mathbf{P}')$.

So in summary, we assume that $p_9 = 0$, and solve for the essential parameters (elements of $\mathbf{P}'$). The eigenvalue decomposition of $\mathbf{P}'^*$ allows us to determine the unknown shift $(p_9 = -\frac{1}{2}\lambda_2)$, and then, $\mathbf{P}$ from:

$$\mathbf{P} = \mathbf{P}' - \frac{1}{2}\lambda_2(\mathbf{P}')\mathbf{I}$$

Finally, the solution of the motion and structure parameters are determined from the equations given earlier in terms of the trace and eigenvectors of $\mathbf{P}^*$ (note that the eigenvectors of $\mathbf{P}^*$ and $\mathbf{P}'^*$ are the same).

7. References

[1] Adiv, G., "Determining 3-D Motion and Structure from Optical Flow Generated by Several Moving Objects," COINS TR 84-07, Computer and Information Science, University of Massachusetts, Amherst, MA, April 1984.

[2] Aloimonos, J., Chou, P. B., "Detection of Surface Orientation and Motion from Texture: 1. The Case of Planes," TR 161, Department of Computer Science, University of Rochester, Rochester, NY, January 1985.

[3] Ballard, D.H., Kimball, O.A., "Rigid Body Motion from Depth and Optical Flow," TR 70, Computer Science Department, University of Rochester, Rochester, NY, 1981.

[4] Bruss, A.R., Horn, B.K.P., "Passive Navigation," *Computer Vision, Graphics, and Image Processing*, Vol. 21, pp. 3–20, 1983.

[5] Buxton, B.F., et al., "3D Solution to the Aperature problem," *Proceedings of the Sixth European Conference on Artificial Intelligence*, pp. 631–640, September 1984.

[6] Fennema, C.L., Thompson W.B., "Velocity Determination in Scenes Containing Several Moving Objects," *Computer Graphics and Image Processing*, 9, pp. 301–315, 1979.

[7] Hildreth, E.C., *The Measurement of Visual Motion*, MIT Press, 1983.

[8] Horn, B.K.P., Schunck, B.G., "Determining Optical Flow," *Artificial Intelligence*, Vol. 17, pp. 185–203, 1981.

[9] Kanatani, K., "Detecting the Motion of a Planar Surface by Line and Surface Integrals," *Computer Vision, Graphics, and Image Processing*, 29, pp. 13–22, 1985.

[10] Longuet-Higgins, H.C., Prazdny, K., "The Interpretation of a Moving Retinal Image," *Proceedings of the Royal Society of London*, Series B, Vol. 208, pp. 385–397, 1980.

[11] Longuet-Higgins, H.C., "A Computer Algorithm for Reconstructing a Scene from Two Projections," *Nature*, Vol. 293, pp. 131–133, 1981.

[12] Longuet-Higgins, H.C., "The Visual Ambiguity of a Moving Plane," *Proceddings of the Royal Society of London*, B 223, pp. 165–175, 1984.

[13] Nagel, H., "On the Derivation of 3D Rigid Point Configurations from Image Sequences," *Proceedings of Pattern Recognition and Image Processing*, Dallas, Texas, 1981.

[14] Negahdaripour, S., Horn, B.K.P., "Determining 3-D Motion of Planar Objects from Brightness Patterns," *Proceedings of Ninth Int. Joint Conf. on A.I.*, pp. 898–901, 1985.

[15] Negahdaripour, S., Horn, B.K.P., "Direct Passive Navigation ," to appear in *IEEE Trans. Pattern Analysis and Machine Intelligence* (also A.I. Memo No. 821, M.I.T. A.I. Lab, Cambridge, MA, February 1985).

[16] Roach, J.W., Aggarwal, J.K., "Determinig the Movement of Objects from a Sequence of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, November 1980.

[17] Subbarao, M., Waxman, A.M., "On the Uniqueness of Image Flow Solutions for Planar Surfaces in Motion," CAR-TR-114, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, MD, April 1985.

[18] Sugihara, K., Sugie, N., "Recovery of Rigid Structure from Orthographically Projected Optical Flow," TR 8304, Department of Information Science, Nagoya University, Nagoya, Japan, October 1983.

[19] Tsai, R.Y., Huang, T.S., Zhu, W.L., "Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch, II: Singular Value Decomposition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-30, No. 4, August 1982.

[20] Tsai, R.Y., Huang, T.S., "Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 1, January 1984.

[21] Waxman, A.M., Ullman, S., "Surface Structure and 3-D Motion from Image Flow: A Kinematic Analysis," CAR-TR-24, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, MD, October 1983.

[22] Waxman, A.M., Wohn, K. "Contour Evolution, Neighborhood Deformation and Global Image Flow: Planar Surfaces in Motion," CAR-TR-58, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, MD, April 1984.